

This is the author's copy of the publication as archived with the DLR's electronic library at <http://elib.dlr.de>. Please consult the original publication for citation.

In-Situ Close-Range Imaging with Plenoptic Cameras

M. Lingenauber; F. A. Fröhlich; U. Krutz; C. Nissler; K. H. Strobl

Copyright Notice

©2019 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Citation Notice

```
@INPROCEEDINGS{lingenauber2019situ,  
  author    = {M. {Lingenauber} and F. A. {Fr\"ohlich} and U. {Krutz} and C. {Nissler} and K. H. {Strobl}},  
  title     = {In-Situ Close-Range Imaging with Plenoptic Cameras},  
  booktitle = {2019 IEEE Aerospace Conference},  
  year      = {2019},  
  pages     = {1--16},  
  month     = mar,  
  doi       = {10.1109/AERO.2019.8741966},  
}
```

In-Situ Close-Range Imaging with Plenoptic Cameras

Martin Lingenauber
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics
Münchener Str. 20
82234 Wessling, Germany
+49-(0)-8153-28-1595
martin.lingenauber@dlr.de

Florian A. Fröhlich
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics
Münchener Str. 20
82234 Wessling, Germany
+49-(0)-8153-28-1487
florian.froehlich@dlr.de

Ulrike Krutz
German Aerospace Center (DLR)
Institute of Optical Sensor Systems
Rutherfordstr. 2
12489 Berlin, Germany
+49-(0)-8153-28-526
ulrike.krutz@dlr.de

Christian Nissler
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics
Münchener Str. 20
82234 Wessling, Germany
+49-(0)-8153-28-3063
christian.nissler@dlr.de

Klaus H. Strobl
German Aerospace Center (DLR)
Institute of Robotics and Mechatronics
Münchener Str. 20
82234 Wessling, Germany
+49-(0)-8153-28-2482
klaus.strobl@dlr.de

Abstract—This paper discusses the concept of plenoptic hand lens imagers for in-situ close-range imaging during planetary exploration missions. Hand lens imagers, such as the Mars Hand Lens Imager on-board the Mars rover Curiosity, are important cameras for in-situ investigations, e.g. of rock layer, minerals or dust. They are also important for the preparation and documentation of other instrument operations and for rover health assessment. Due to the small working distance between object and the camera's main lens, significant physical limitations affect the imaging performance. Most evident is the limited depth of field of a few millimeters for working distances of a few centimeters. This requires a highly accurate positioning of the camera and also limits the in-focus content of an image significantly. Hence, in order to have an extended object completely in focus, a sequence of images, each being focused to a different distance, is required. A single, passive camera is insufficient to compute depth from a single shot; only the combination of multiple images, either taken from different vantage points or at different focal settings, allows this. To overcome those limitations, we propose the use of plenoptic cameras as hand lens imagers. From a single exposure, they allow to create an extended depth of field image and at the same time a metric depth map while maintaining a more open aperture. These and other advantages might make it possible to omit space grade focus mechanisms in the future. A plenoptic camera is achieved by adding an additional matrix of lenslets shortly in front of the image sensor of a conventional camera. Hence, available space camera hardware can be used to form a new type of sensor. Due to its recording concept, a plenoptic camera maintains the depth of the scene as it is projected into the camera by the main lens. Thanks to the parallax between the lenslets, it is possible to compute depth via triangulation for each image point as well as a high resolution 2-D extended depth of field image. This paper provides an overview of the state of the art of hand lens imaging from which we derive a set of common requirements for future devices. We briefly introduce the plenoptic camera technology and provide first experimental results on the imaging performance based on samples of test targets and rocks. The results show that our preliminary plenoptic camera setup can comply with the requirements for in-situ hand lens imaging in terms of image quality, depth estimation and the usability for planetology.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
2. STATE OF THE ART OF HAND LENS IMAGING	2
3. PLENOPTIC CAMERA CONCEPT	5
4. EXPERIMENTS	8
5. RESULTS	9
6. DISCUSSION	13
7. CONCLUSION	13
ACKNOWLEDGMENTS	14
REFERENCES	14
BIOGRAPHY	16

1. INTRODUCTION

This paper proposes to use plenoptic cameras for the next generation of hand lens imagers for in-situ micro-imaging during planetary exploration. A Hand Lens Imager (HLI) is the remote version of a geologist's hand lens. As such its main purpose is to provide images of small geological features, such as cracks, layers or single grains of dust. A HLI can serve as a camera on-board a rover or lander that provides in-situ micro-imaging with resolutions of a few $\mu\text{m}/\text{px}$. Examples of HLIs are the Mars Hand Lens Imager (MAHLI) aboard the Mars rover Curiosity [1], and the Close Up Imager (CLUPI) in preparation for the ExoMars rover [2] planned to be launched in 2020. In addition to images of single features of scientific interest, HLIs are often used to provide context information for the preparation and documentation of other instruments' sampling activities. And they also turned out to be useful for rover operations, health assessment or for taking landscape views [3], making them valuable instruments besides their scientific purpose.

In order to provide the maximum resolution of $14.5 \mu\text{m}/\text{px}$ for MAHLI [1] or $7 \mu\text{m}/\text{px}$ for CLUPI [2], the cameras are operated at small Working Distances (WD) of approximately 2 cm and 10 cm, respectively. Here, the WD is the distance

between front surface of the main lens and the object. This results in a shallow Depth Of Field (DOF) of approximately 1 mm for MAHLI as well as for CLUPI [2, 4]. The limited range in which an object is in focus is often not sufficient for scientific purposes and poses high requirements on the exact positioning of the camera in front of a target. In order to overcome this limitation, MAHLI and CLUPI each use a complex focus mechanism that allows to refocus the camera in small steps, thus enabling it to provide an Extended Depth Of Field (EDOF). They record image sequences of up to 8 or 16 images, respectively, and process them on-board in order to produce an EDOF image and as a byproduct also a depth map as shown in Figure 2.

The use of conventional camera technology for HLIs increases the complexity, volume, mass and risk of failure of the system due to moving parts in a space environment. Therefore, we propose to use plenoptic cameras for future HLIs to overcome these limitations. Plenoptic cameras inherently possess multiple advantages over conventional cameras that become most evident at small WDs. Two of the key advantages are that with a single plenoptic recording, it is possible to extend the DOF multiple times and to create a depth map [5]. This requires no moving parts as plenoptic cameras allow to shift certain operations from hardware to software, which in turn enables a more robust design and renders them a good choice for the harsh space environment. It also means that plenoptic cameras can be re-calibrated during the main mission, by using four to eight images of a known calibration target [6]. A plenoptic camera can be created by modifying a conventional camera with a single optical component, the so called Micro-Lens Array (MLA). It is mounted with a gap of a few micrometers in front of the image sensor (cf. Section 3). Thus, it is possible to use space qualified camera technology, sensors and optics in order to create a new type of image sensor. The remaining challenge in establishing this technology is the development of space grade processing of plenoptic image data.

The concept of using plenoptic cameras for in-situ plane-tology has been shown in [7] along with qualitative results. Hence, in order to extend the work of [7], the main contributions of this paper are

- the first extended discussion on the use of plenoptic cameras for in-situ hand lens imaging
- the first quantitative experimental results on the 2-D resolution and the depth measurement capabilities as well as examples that show the usability for planetology
- a list of common objectives and requirements for the development of future HLI systems

2. STATE OF THE ART OF HAND LENS IMAGING

Three HLIs that represent the current state of the art are used in this work as reference camera systems. We use their performance values to compare our results with. Additionally, they are used to derive a common set of goals and high level requirements that we propose to use as a basic guideline for future HLIs, both plenoptic and conventional camera based ones. Namely, these are MAHLI aboard the Mars rover Curiosity [8], CLUPI [2] in production for the ExoMars 2020 rover, and MasCam, the camera aboard the recently landed asteroid lander MASCOT which is part of the Hayabusa 2 mission to the asteroid Ryugu [9]. Legacy devices such as the Micro Imagers (MI) aboard the Mars Exploration Rovers

(MER) [10], the Robotic Arm Camera (RAC) aboard the Mars lander Phoenix [11, 12] or ROLIS on the comet lander Philae [13, 14], are not further considered because MAHLI and MasCam use heritage of some or one of them. Due to the lack of information, this work does also not consider the context imager which is a component of the PIXL instrument for NASA's Mars 2020 rover [15].

First, we provide a set of common goals and requirements compiled from available information on those systems and from discussions with DLR planetary scientists. This set can serve as a baseline for the development of future HLIs. Due to the different instrument designs of the HLIs and due to different missions, target bodies and mounting positions on the hosting space crafts, the provided goals and requirements are rather high level. Second, we provide a brief description of the reference HLIs and examples of data products created with them in order to summarize the current state of the art. An overview of the main specification values of the reference HLIs is presented in Table 1.

Common goals

From a planetary scientists point of view, a HLI shall work as a remote hand lens for geological and astrobiological research. As such it shall provide high resolution color images in order to observe targets and sites of interest in-situ from meter to sub-millimeter range. It shall help to characterize geology, mineralogy and physical properties of objects and surfaces such as rocks, dust, sediments, regolith and even the landscape. Thus, images that allow to investigate items such as texture, shape, structure, morphology, layering or particle size distribution of a sample area are required.

A HLI should also act as a complementary instrument that provides context on a sample site with images taken from different distances and vantage points. These images can also be combined to provide mosaics or panorama views. This can support the operation of other instruments in terms of selection of sample areas, preparation of operation procedures and, documentation of the sampling. The high resolution and the operation close to a target shall close the resolution gap between microscopes, if available, and instruments operated from a larger distances, e.g. the panorama cameras on a rover's mast or the cameras of an asteroid orbiter. Depending on the mounting position and the possibility to move the HLI, it should also support engineering tasks, tests and monitoring of the spacecraft and its components.

From an operations point of view, the HLI should work in a semi-autonomous fashion due to long signal runtimes. On-board pre-processing and compression are required in order to reduce the amount of data to be transmitted. Experience from the Curiosity mission shows that a larger Field Of View (FOV) reduces the amount of images required for several tasks, which results in a more efficient operation and a higher probability of instrument use [3].

Specific mission goals can guide the HLI design more towards a spot sampling imager, such as MAHLI or CLUPI, or one with a larger FOV in order to provide more surveying opportunities as it is done with MasCam.

Common requirements

2-D particle resolution depends on the mission goals. It shall allow to resolve particles of at least 1 mm in size. Investigations on sediments require a resolution of at least 62 μm in order to distinguish between sand and silt according

to the Wentworth grain size scale [16]. Investigations on the morphology of grains or astrobiology may require a higher resolution.

2-D ground resolution is the per object point resolution of the HLI system in $\mu\text{m}/\text{px}$ and, as a rule of thumb, it shall beat least four times better than the required particle resolution.

DOF should be as large as possible, in order to reduce the number of images that are required to provide a complete in-focus image of a target. Due to the high variability of possible sample sites, it is difficult to provide a finite value, but we aim for a DOF of at least 1 cm, and preferably 5 cm at the minimum WD.

Depth is required to complement the 2-D images as it provides additional context. Depth measurements shall provide a dense depth map over the complete FOV.

Depth range is the range from the target surface towards the camera in which depth can be determined. It should at least be equal to the DOF, thus 1 cm to 5 cm.

Depth ground resolution is the finest depth difference that can be measured between two object points, thus it allows to resolve particles and their shape in depth. It is a relative measurement between two points within the depth range. For a triangulation based depth estimation, it should be approximately four times less than the 2-D ground resolution.

Depth precision is the absolute measured distance between the HLI and the target. It is regarded for this work to be less important for scientific measurements than for engineering tasks, such as distance checks for the camera positioning or for hazard avoidance. It should be at least $1/3$ of the DOF at minimum WD in order to allow for a reliable and save positioning of the camera such that the target is at least $2/3$ in focus in the depth direction.

FOV of the camera system should provide a ground resolution cell at minimum WD of at least $2\text{ cm} \times 2\text{ cm}$ in size in order to provide enough area to determine the particle size distribution of a sample.

Image and depth map size should be at least 1 Mpx, preferably higher in order to increase the amount of details in a single image.

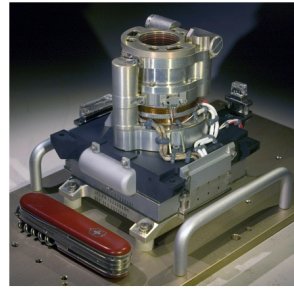
Working distance is the distance between the objective's front face and the target. It depends on the required 2-D and depth resolution, DOF and FOV. The WD should be as large as possible in order to ease the operation and to reduce the risk of collision of the camera with a target.

Color is required to provide additional context, allow a preliminary distinguishing of materials but can also aid to distinguish between adjacent particles [3]. A color calibration of the images is required in order to account for the different illumination conditions on another planet. This requires a color calibration target for in-situ calibration.

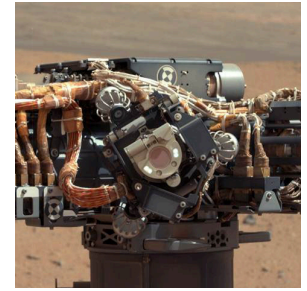
In-focus images are required from minimum WD to infinity in order to support different operations up to landscape views. The focus must not be continuously available in the complete range. Consequently, the two HLIs for Mars use a focus mechanism to comply with this requirement.

On-board processing for the semi-automatic operation must at least provide an autofocus or the detection of the ideal focus position for the positioning of the camera. Processing of the raw data to create the final data products, e.g. 2-D EDOF images or depth maps or High Dynamic Range (HDR) images, and their compression prior to transmission is also required in order to reduce the amount of data that needs to be transmitted.

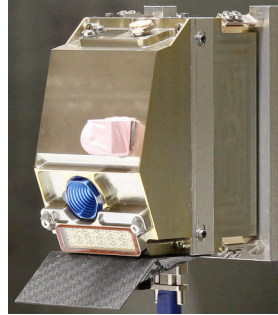
Illumination is required to allow recording at day and night or to minimize the effect of shadowing, e.g. due to other parts of the spacecraft. It can also increase the information content of images or can also be another way of creating color images and therefor of distinguishing materials as with the



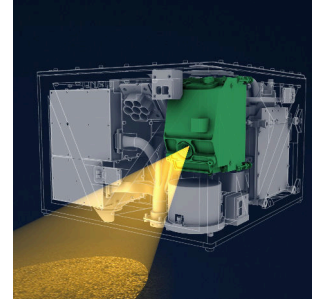
(a) MAHLI camera head.
Image credit: NASA/JPL-Caltech/MSSS



(b) MAHLI mounting position on robotic arm turret.
Image credit: NASA/JPL-Caltech/MSSS



(c) MasCam. Image Credit: DLR



(d) MasCam mounting position in MASCOT Lander.
Image Credit: DLR

Figure 1: Photos and artist impressions of reference HLIs

multi-color illumination of MASCOT.

Reference HLI description

MAHLI [8] is a 2 Mpx color camera. It consists of a camera head shown in Figure 1a, a digital electronics assembly, and a calibration target. The camera head is mounted on the turret at the end of the robotic arm of Curiosity (cf. Figure 1b), the digital electronics assembly is inside the rover body and the calibration target is on the chassis. The camera design explicitly considers the extended camera mobility due to its mounting on the robotic arm. Hence, as the camera can be positioned relatively freely to a target and in order to increase the context provided by an image, MAHLI has a rather wide FOV of 34° to 38.5° (diagonal). It results in a minimum WD of 2.1 cm to achieve the maximum 2-D ground resolution of $14\mu\text{m}/\text{px}$ with a minimum DOF of 1 mm. The camera is equipped with a focus mechanism and an autofocus algorithm, which allows the camera to focus from minimum WD to infinity. On-board processing allows to process an image sequences of up to 8 images to create an EDOF image via focus stacking, a depth map, and different compression levels of the data in order to minimize the amount of downlink data. The focus mechanism employs a stepper motor with a system of gears to move the lens focus group and the protective front cover [17]. The MAHLI team experimentally determined a calibration function that relates stepper motor count to working distance in object space [18]. The complex focus mechanism required a re-design during the development due to issues with the mechanisms and the lubrication [17]. An illumination system of two pairs of white LEDs and two UV-LEDs at the camera's front face provides day and night recording capabilities.

CLUPI [2] is a 4.7 Mpx color camera system mounted to

the drill box of the ExoMars rover. It consists of the camera head which includes the electronics and processing parts and a separated calibration target that is mounted to the rover's chassis. Despite the drill box's limitation to two degrees of freedom (rotation around one axis and up-down translation), CLUPI can provide multiple different views and therefore also high imaging flexibility by using multiple mirrors that are mounted to the rover's chassis. By this, it is possible to take images of sample sites, the drill as well as the drill hole, and of the landscape. The FOV is $(14 \pm 2)^\circ$ (diagonal), resulting in a maximum resolution of $7 \mu\text{m}/\text{px}$ with a minimum DOF of 1 mm at the smallest WD of 10 cm. The camera also comes with a focus mechanism, based on a flexible structure which does not require lubrication [2]. As more information about the mechanism is not available at the time of writing, an evaluation of its complexity is difficult. The on-board processing of CLUPI provides an autofocus and an autoexposure algorithm. It allows to create EDOF images and depth maps from focus stacks of up to 16 images, binning and Region Of Interest (ROI) recording as well as compression for data transmission savings.

MasCam [9] is a 1 Mpx panchromatic camera mounted inside the shoe-box sized android lander MASCOT. Due to mission specific objectives and requirements, it follows slightly different design principles than MAHLI and CLUPI. Being the camera of one of the first mobile lander on an asteroid, the goal is more to provide an overview of the landscape and in-situ context for MASCOT's as well as Hayabusa 2's instruments, than to achieve the highest resolution for spot samples opposed to the other HLIs. Additionally, the design is largely constrained by the small dimensions of the spacecraft and the camera being fixed inside the structure of MASCOT. A dark image calibration target is available during the cruise flight aboard Hayabusa 2. The camera comes with a wide FOV of 72.5° (diagonal) and is designed to provide a DOF from 150 mm to the local horizon, i.e. infinity. This is achieved by reducing the aperture to the diffraction limit, and by tilting the sensor by some degrees with respect to the objective's optical axis, resulting in a Scheimpflug configuration. On the one hand, this large DOF and the wide FOV reduce the ground resolution to $150 \mu\text{m}/\text{px}$. However, on the other hand, no focus mechanism is needed. This also makes it difficult if not impossible to compute depth maps when the lander is not moving. A LED array provides illumination to the near field at night with four different colors: red, blue, green channels in the visible range and one channel in the near infrared. Thus, MasCam provides panchromatic images at day and multi-color images at night. On-board processing provides image statistics, autoexposure, HDR images from three pictures and image compression.

Data products from HLIs

A selection of possible data products created with HLIs are shown in Figure 2 and Figure 3. These examples show the bandwidth of usage of HLIs, i.e. from scientific imaging in Figure 2, to context imaging in Figure 3a and Figure 3d to imaging for engineering purposes as in Figure 3b.

The first and last image and the resulting focus stacking products of a sequence of images recorded with MAHLI on Sol 998 are shown in Figure 2, which is also an example of the main scientific use of MAHLI. As can be seen, the small WD for this sample resulted in a shallow DOF as there is significant blur in the foreground or background, respectively, of the shown images. The EDOF image in Figure 2c was created by merging the eight images in a focus stack and it shows a high resolution and image quality and

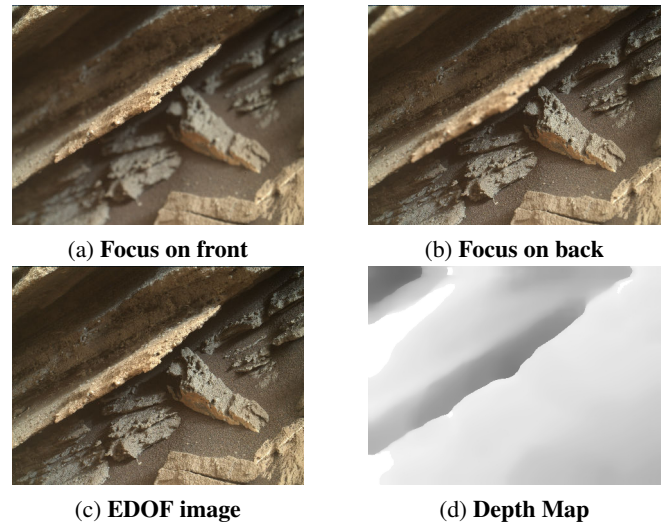


Figure 2: Images from a sequence made by MAHLI on Sol 998 and the generated EDOF image and depth map. Eight images were merged for this focus stack. Image credit: NASA/JPL-Caltech/MSSS

the whole scene is in focus. The depth map on the other hand shows the limitation of the focus stack approach for depth measurements, as it provides only a low level of detail and low depth resolution. This makes it difficult to detect finer details such as the grains in the clast. As the digital numbers in the depth map are related to the focus position of the merged input images, only eight known distances contributed to the depth map and other values inbetween are interpolated, which makes it a rudimentary range measurement for arm positioning and coarse context information [18].

Figure 3a shows a combination of three images taken with MAHLI on Sol 47 of a target rock named *Jake Matijevic* and thus the typical operation mode [3]. The collage in Figure 3a contains images taken at a WD of 25 cm, 5 cm and 2.5 cm which gives an impression of different ground resolution cells and the resolution differences at these WDs. With this mode of operation MAHLI first provides a wider FOV for more context, second an image with nearly the same resolution as the MIs aboard Spirit and Opportunity for comparability and third the intended full resolution image.

An image by MAHLI recorded on Sol 411, probably for wheel inspection and thus an engineering use case, is shown in Figure 3b. At this larger WD, the DOF is wide enough for this use case and thanks to the high resolution of MAHLI a sufficient amount of details are visible, e.g. for damage assessment.

A section of one of the highest resolved images taken with MAHLI during the primary mission on Sol 411 is shown in Figure 3c. The following values and explanation are a summary from [18]. Here, the coin on MAHLI's calibration target was recorded at a WD of 3.7 cm and the image shows single grains of accumulated dust. In this image, grains down to a size of $28 \mu\text{m}$ to $42 \mu\text{m}$ were identified, thus allowing to distinguish clearly between silt (grain diameter $\leq 62.5 \mu\text{m}$) and very fine sand. To get a feeling for the dimensions in the image, the coin itself has a diameter of 19 mm and the small pile of grains visible in the lower left corner of Figure 3c is approximately $540 \mu\text{m}$ long in the vertical direction of the image. Due to the solar incidence angle of about 71.3°

Table 1: Specification of reference HLIs

Value	MAHLI [8]	CLUPI [2]	MasCam [9]
sensor	CCD, RGB	CMOS, RGB	CMOS, mono
pixel size	7.4 μm	7.8 μm	15 μm
image size	1600 px \times 1200 px	2652 px \times 1768 px	1024 px \times 1024 px
focus mechanism	stepper motor	flexible structure	fixed
mass	578 g (camera head)	932 g (total)	407 g (total)
min. WD	2.1 cm	10 cm	15 cm
FOV diagonal	33.8° to 38.5°	(14 \pm 2)°	72.5°
IFOV	402 μrad	N/A	1014 μrad
focal length	18.4 mm	N/A	14.8 mm
f-number	9.8	N/A	16
min. DOF	1 mm	1.2 mm	150 mm to ∞
examples of ground	14 @ 2.1	7 @ 10	150 @ 15
resolution in	31 @ 6.6	39 @ 50	
$\mu\text{m}/\text{px}$ @ WD in cm		79 @ 100	
examples of ground	2.24 \times 1.68 @ 2.1	1.9 \times 1.3 @ 10	15.36 \times 15.36 @ 15
resolution cell in	49.6 \times 37.2 @ 6.6	10 \times 7 @ 50	
cm \times cm @ WD in cm		21 \times 14 @ 100	

during the recording, it was also possible to use the shadows casted by the single grains for size determination.

Figure 3d shows an image recorded during the landing approach of MASCOT on Ryugu. The image was taken several meters above the surface, shortly after the first contact of the lander. This is an example of using a HLI for imaging a larger context and as a descent imager, one of the goals of MasCam [9].

3. PLENOPTIC CAMERA CONCEPT

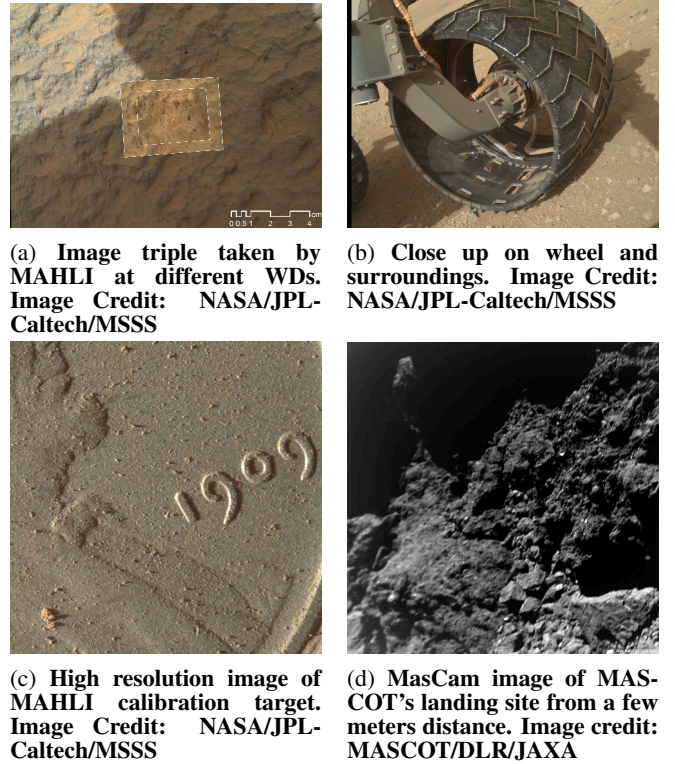
The plenoptic function is a theoretical concept with the aim to describe the available information for an observer in the form of light rays at any point in time and space with multiple dimensions [19]. An ideal observer with a FOV of 360° at position V_x, V_y, V_z in space can receive light rays from all directions θ, ϕ for all wavelengths λ at all moments in time t , which results in a measurement of the plenoptic function

$$P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z). \quad (1)$$

When a static scene and single wavelength is assumed, the plenoptic function reduces to 5-D as t and λ become constant. This can be further reduced to 4-D when the scene is considered to be inside a boundary box and that the intensity does not change along the light ray's direction of propagation, hence V_z becomes constant. This results in the 4-D light field introduced by Levoy and Hanrahan [20], as well as the similar concept of a lumigraph introduced by Gortler et al. [21]. Now, the position and orientation of a light ray inside the bounding box can be described by its two intersection points at u, v and s, t on the parallel planes resulting in

$$L = P_L = (u, v, s, t) \quad (2)$$

which is a description of the recorded light intensity or radiance in the light field. Intuitively, this two-plane parametrization of a light field can be explained by uniformly placing a collection of pinhole cameras on the s, t -plane viewing

**Figure 3: Further examples of images taken with HLIs**

towards the u, v -plane. This allows to visualize the light field such that each of the cameras at a position s, t would view all points on the u, v -plane from a slightly different vantage point, thus creating $s \times t$ unique views of the $u \times v$ points on the u, v -plane. The s, t -plane could also be a matrix of pin-holes inside a conventional camera. If the matrix is located at a short distance b in front of the sensor and with a known distance a behind the u, v -plane (seen from the main

lens), the most basic and idealized form of a plenoptic camera would be achieved. Here, each pixel within the FOV of a s, t -pinhole records the intensity of a light ray from a different direction, thus u, v -coordinate, which results in a recording of the light field in Equation 2.

By considering the above, the difference to a conventional camera becomes more clear. Its main lens is focused on the sensor plane on the image side of the lens in order to produce a sharp image. Hence, when the main lens is focused on a 3-D object, it images it as a 2-D image on the sensor plane. Therefore, all light rays that leave an object point are converged by the main lens to an image point and their intensity is summed up, which results in a loss of information. In contrast, a plenoptic camera splits up the pencil of rays once more at the s, t -plane and thus enables to record the intensity of light rays from different directions on different pixels. By this and given the known geometry of the described optical setup, a 2-D sensor can record a 4-D light field. Hence, a camera which records the subset of the plenoptic function described in Equation 2 is also known as a light field camera. This recording concept makes it also necessary to process the raw data to create a conventional 2-D image. However, as the direction of the recorded light is encoded in the pixel position and the geometrical relationship between the s, t and u, v -plane, the light field also contains depth information.

Plenoptic camera concepts

In order to record the light field in practice with a single camera, Ng et al. [22] proposed to modify a conventional digital camera by placing an MLA shortly in front of the camera sensor. The s, t -plane is now represented by the MLA plane and the u, v -plane is located at the main lens principle plane. Through a micro-lens, each sensor pixel with its finite FOV sees a small excerpt of the main lens plane, which represents a single sample of the u, v -plane. On the other side, each micro-lens is regarded as a single sample of the s, t -plane as all light rays within its FOV go through its principle point. As such, a micro-lens images only the part of the main lens aperture within its FOV. In order to achieve this imaging process, the main lens is focused onto the MLA plane and the micro-lenses, due to their relatively small size in comparison with the main lens, are considered to be unfocused i.e. with a focus to infinity. Hence, this plenoptic camera concept is known as the Unfocused Plenoptic Camera (UPC), or standard plenoptic camera, or plenoptic camera 1.0. Due to the finite size of the micro-lenses and the pixel's FOV, this results in a box-filtered sampling of $L(u, v, s, t)$ [22]. For this design, the maximum spatial resolution, i.e. the resolution of the rendered 2-D image equals the number of micro-lenses in the MLA and the angular resolution is equal to the number of pixels under a micro lens [23].

A different plenoptic camera concept, proposed by Lumsdaine and Georgiev, is known as the Focused Plenoptic Camera (FPC) [24, 25]. Here, the main lens is focused in front or behind the MLA and the micro-lenses are focused on the image that is projected by the main lens into the camera. This design allows a trade-off between spatial and angular resolution of the light field and thus can achieve a higher resolution of the finally rendered 2-D images than the UPC.

A further improvement of the FPC is the Multi-Focus Plenoptic Camera (MFPC), for which the MLA contains micro-lenses with different focal lengths [5, 26]. By choosing the focal lengths such that the DOFs of the different micro-lens types just touch, it is possible to increase the DOF of the

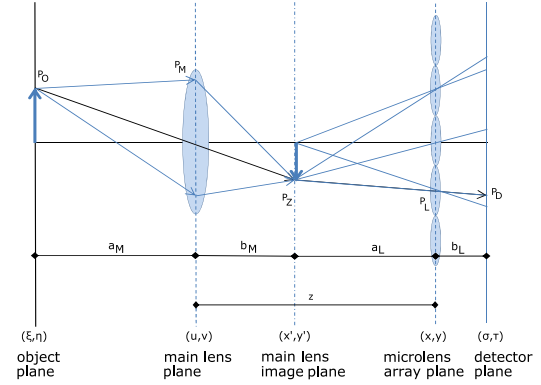


Figure 4: Focused plenoptic camera

MLA significantly. This increases the DOF of the MFPC in comparison to a FPC or UPC and even more if compared to a conventional camera, especially at small WDs [5]. In practice, the spatial resolution, i.e. the final 2-D image resolution, of MFPCs is approximately a quarter of the sensor resolution, which is a more acceptable reduction of the spatial resolution than for UPCs.

The different sampling of UPC and FPC and the resulting resolution differences are not only caused by the different focusing of the main lens, but also from a different distance b_L between the MLA and the sensor plane. In case of the UPC, this distance is exactly the focal length f_L of the micro-lenses. Whereas in the case of the FPC or MFPC $b_L \neq f$ and b_L is adjusted in order to achieve the desired trade-off between spatial and angular resolution [23].

Important for all of the described concepts is the f-number matching rule, i.e. the effective f-number k_{eff} of the micro-lens and of the main lens have to match in order to achieve the maximum possible size of a micro-image without causing an overlap between adjacent micro-lenses [5, 22]. The effective f-number k_{eff} is relevant for objects at close proximity to the camera and is defined as

$$k_{eff} = k_{\infty}(1 + |m|) = 1/2NA_i \quad (3)$$

where k_{∞} is the photographic f-number based on the assumption that the object is relatively far away in comparison to the focal length [27, 28]. The photographic f-number can usually be found on the lens housing. The magnification of the objective is given by m , here as the absolute value. NA_i is the image side numerical aperture of the objective.

Besides the shown MLA based plenoptic camera concepts, other methods are known to record light fields or sample parts of the plenoptic function and are discussed in [29] and [30]. So far, only MLA based plenoptic cameras are beyond a prototype phase and commercially available. Due to this fact and the noted advantage of higher lateral resolution, we consider the MFPC concept the best choice amongst the different plenoptic camera concepts for in-situ micro-imaging and thus focus on it during our analysis.

Optical model

Ray tracing gives a first order estimation of the optical system of a light field camera. Though not taking diffraction or aberration into account, this approach is convenient for extracting information about the footprints that a light ray

makes by passing a system given in Figure 4. There is a main lens plane (u,v), a micro-lens array plane (x,y), and a detector plane (σ, τ). The focal lengths f_M of the main lens and f_L of the micro-lenses are arbitrary, though dependent from each other due to the f-number matching rule [22]. We assume $z \gg b_L$. The object plane is given by (ξ, η) . In addition, P_O is a point in the object plane, P_Z a point in the virtual plane (x', y'), and P_D the point in the detector plane to which the micro-lens projects P_Z from the virtual image plane. Let the center of the main lens plane be the origin of the coordinate system. The object plane point $P_O(\xi, \eta, -a_M)$ is mapped to a detector plane point $P_D^i(\sigma^i, \tau^i, b_M + b_L + a_L)$ in the following way:

$$\begin{aligned} \xi &\mapsto \sigma^i = -\frac{\xi f_M b_L}{a_M f_M - z(a_M - f_M)} + \frac{x_{off}^i(f_M - a_M)b_L}{a_M f_M - z(a_M - f_M)} \\ \eta &\mapsto \tau^i = -\frac{\eta f_M b_L}{a_M f_M - z(a_M - f_M)} + \frac{y_{off}^i(f_M - a_M)b_L}{a_M f_M - z(a_M - f_M)} \\ -a_M &\mapsto b_M + b_L + a_L \end{aligned} \quad (4)$$

Here, i is the certain micro-lens the ray passes before hitting the detector. For detailed information and the viceversa projection see [31]. As mentioned above, the geometrical approach is sufficient to get a first impression of how a pleoptic system works. However, it is not exact. Due to ever smaller pixel sizes used in detectors nowadays combined with low aperture and the small sized micro-lenses, diffraction effects are not negligible. Therefore a wave-theoretical approach is necessary. The Point Spread Function (PSF), which has an effect on the resolution, is an essential parameter describing the performance of the camera. The wave-optical theory has been outlined in [31] and shows a significant different footprint of the light on the sensor, caused by diffraction on the optical path, in comparison to the more simple geometric model presented here.

Plenoptic data processing

The optical model mentioned in the previous section is the least required knowledge for processing plenoptic raw data in depth and 2-D data. In order to achieve the required parameters of the model a calibration is required as an initial step. First, the so called MLA calibration provides parameters such as the lateral MLA to sensor offset, the micro-lens size for the given focus setting and a grey image, all necessary to create depth maps and 2-D images. Depth values created only with a MLA calibration are provided in virtual depth [5], a value that reflects the triangulation geometry required for its measurement but it has no metric scale.

In order to produce metric depth maps and 2-D images with a metric relationship to object space, the values for the MLA to main lens distance z and the MLA to sensor distance b_L need to be calibrated [6, 32]. As with conventional cameras also the parameters for the camera intrinsics and for the optical distortion need to be estimated to achieve a high precision. Several metric calibration approaches for MFPCs are reported in literature [6, 32–34] that require several images of a calibration target with known geometry. Most often a rigid plate with e.g. a checkerboard pattern or a point pattern is recorded in different poses. For this work, the approach by [32] was applied as it is part of the currently available standard plenoptic processing pipeline for MFPCs. Already with the MLA calibration, the processing pipeline

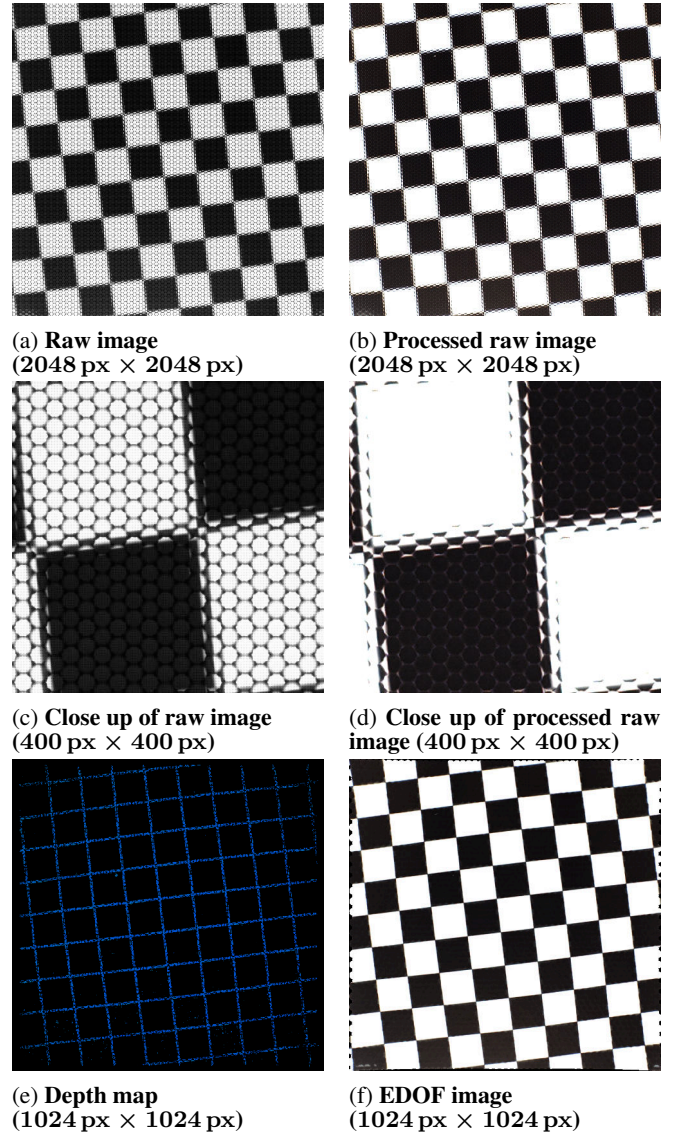


Figure 5: Plenoptic processing pipeline

described in [5], and exemplarily shown in Figure 5 with a checkerboard test pattern as an object, can be used. First, the raw image in Figure 5a is pre-processed, e.g. for brightness equalization and the detection of micro-images, resulting in the processed raw image shown in Figure 5b. Second, a depth estimation is performed resulting in a depth map shown in Figure 5e followed by the combination of corresponding pixel intensities in order to form a 2-D EDOF image shown in Figure 5f.

The depth estimation is a triangulation based on correspondences found in adjacent or near micro-lens images [5]. Those correspondences can for example be established by using the Sum of Absolute Differences (SAD) as proposed in [5]. Therefore, the depth estimation is limited primarily to well textured areas or edges as the correspondence search requires sufficient contrast. Hence, the depth estimation becomes a multi-view stereo vision problem, as each micro-lens acts as a small camera. This is a computational expensive step, but it can be well parallelized. The depth estimation is required prior to the 2-D image generation in order to choose the correct micro-lens images that result in the maximum

resolution but also as otherwise artefacts, especially at object boundaries, are the result [5, 35]. In order to create a 2-D image point, all micro-lenses that can see the same point of the main lens image are chosen based on depth information. The pixel position on which the main lens point is projected, is computed and the final 2-D intensity value is a weighted average of corresponding pixel values under the previously selected micro-lenses [5].

Due to the MLA geometry of an MFPC with up to three different micro-lens types and the increased DOF, the final 2-D image has a DOF multiple times larger than the one of a conventional camera. At small WDs, the DOF was reported to be up to 6 times larger than the one of a comparable conventional camera [5]. In other words, an MFPC achieves a significant DOF extension due to its recording concept and the appropriate data processing. Thus, in order to achieve an EDOF image an MFPC does not require a focus mechanism because hardware operations are shifted to software. The final size of the 2-D images and depth maps is approximately a quarter of the sensor size, because lateral resolution is converted into depth information.

4. EXPERIMENTS

The experiments should provide the first real measurements to allow a preliminary evaluation of the plenoptic camera concept for in-situ micro-imaging based on real data and not only on theoretical and conceptual findings. For this we chose the requirements which we regard as the design drivers at the current state of the technology. These are the 2-D resolution, the DOF and the depth measurement capabilities. For the latter we determined the depth resolution, depth precision and depth noise. At the time of writing, our experiment setup did not allow a determination of the DOF, yet. Thus this is left for future work. Please note, that our plenoptic camera system under test is not yet optimized for the requirements. However, it uses a sensor of which a space qualified version exists. Nevertheless, we argue that from this preliminary test setup it is possible to see if the conceptional advantage are evident in practice. At the moment we provide the tests for a single WD of 160 mm in order to gain the maximum resolution. In the future we plan to perform our measurements at multiple WDs.

Setup

We employed a Raytrix R5-C color MFPC equipped with a Fujinon HS35SA-1 1:1.4 35 mm main lens for our experiments. This camera uses a CMV4000 CMOS sensor with $2048 \text{ px} \times 2048 \text{ px}$ with a pixel pitch of $5.5 \mu\text{m}$. The resulting image and depth map size is $1024 \text{ px} \times 1024 \text{ px}$, which is the lowest image size stated in the requirements in Section 2. The main lens was set to the nearest possible focal setting and the aperture was set to an effective f-number of approximately 2.4 in order to comply with the f-number matching rule. The camera setup provided a FOV of approximately 58 mm at a WD of 160 mm. At this WD, the highest lateral resolution in combination with the measurement of depth could be achieved. As shown in Figure 6 the camera was mounted on a vertical linear axis that allowed to accurately set the WD with respect to an object. Objects were placed on a XY-table below the linear axis in order to precisely and repeatably position them in lateral direction within the FOV. Test charts and objects that required a 45° tilt with respect to the optical axis were mounted on a tilted board which was also mounted on the XY-table as shown in Figure 6. During the experiments, the following test objects were used:

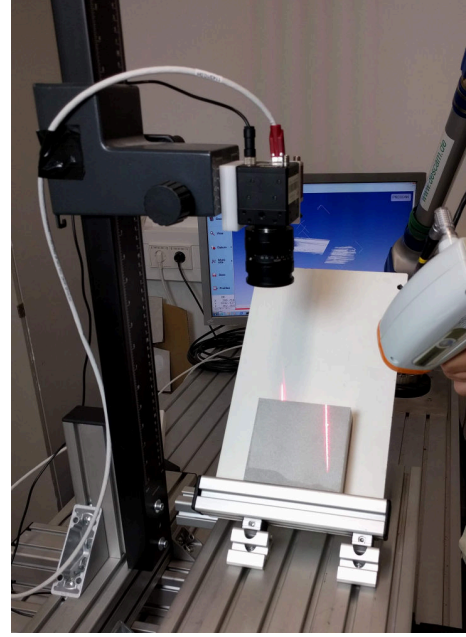


Figure 6: Experimental setup during laser scanning

- Metal block tilted by 45° with a sandblasted surface that provides a fine texture as it is needed for triangulation based depth measurements. This allows to determine the depth measurement quality for a known plane over a larger depth range.
- Checkerboard pattern calibration chart with size DIN-A4 and an edge length of the squares of 6 mm
- Point pattern calibration chart whose points have a diameter of 2 mm
- Lensgrid test chart generated with MTF-Mapper, scaled to DIN-A8
- Basaltic rock with fine structures, pores and large depth differences to show the possible depth range and the resolution of the camera with a geological object

To generate ground truth data for the depth measurements, we used a high precision metrology arm that was equipped with a laser scanner as shown in Figure 6. This allowed us to measure the position of objects, to determine the exact tilt angle of planes and to create a reference 3-D model of the object for the depth experiment. The software RxLive 4.0.50 by Raytrix was used for recording and processing the plenoptic data. With its help the camera was calibrated and 2-D images, EDOF images, depth maps and point clouds were created for further analysis. In order to determine the 2-D resolution of the plenoptic camera system, we measured the Modular Transfer Function (MTF) with help of the open source software MTF-Mapper in version 0.7.5 [36]. The software allowed to create test charts with slanted edges for the determination of the resolution.

Methods

Calibration—Besides the necessary MLA calibration, we performed the metric calibration approach provided by the RxLive software, which was the reference implementation of [32]. Here a sequence of 21 images of a point calibration pattern was necessary for the calibration to converge to an acceptable quality level.

2-D resolution—For the determination of the 2-D resolution, we used the slanted edge method [37]. This measured the Spatial Frequency Response (SFR) of the system, which can be converted into the MTF in terms of image space resolution versus contrast. In other words, by measuring how sharply an edge is imaged by the camera system, the frequencies in an image can be determined, which in the end represent the system's spatial resolution and the PSF (cf. Section 3). By using edges distributed all over the FOV and with different rotations, it is possible to determine a potential resolution loss towards the image boundaries and to determine the SFR in all directions. It is important to note, that as an image taken with the camera is used for the measurement, the slanted edge method determines the system MTF and not the MTF of only the optics. Hence, for a conventional camera it would be the combined MTF of the lens, sensor and the camera read-out electronics. In case of a plenoptic camera, the MTF measurement is performed on a 2-D or EDOF image generated with the plenoptic processing pipeline as explained in Section 3. Thus, in contrast to a conventional camera, the measured system MTF additionally contains the SFR of the plenoptic data processing as part of the camera system. Additionally, as said previously, the processing combines values recorded by multiple pixels and results in an image resolution of a quarter of the sensor resolution. Therefore, we assume twice the pixel pitch of the sensor for the MTF measurement, i.e. in this case a pixel pitch of 11 μm .

The test was performed with MTF-Mapper and with a MTF-Mapper generated lensgrid test chart (cf. Figure 7). In order to fit the complete or at least a large part of the test chart in the FOV, we scaled the lensgrid test pattern to DIN-A8 format as seen in Figure 7. Due to the small size of the test chart, the circular fiducial patterns required to compute the orientation of the chart with respect to the camera, could not be used. Hence, we tried to orientate the camera as parallel to the test chart as possible in order to reduce the standard deviation of the MTF measurements to a minimum (cf. Figure 8). Due to the small WD and the resulting small scale of the test chart, we assume that even small rotational offsets result in a larger deviation of the MTF measurements compared to recording at a larger WD. The test chart was printed on thick paper with a weight of 350 g/m^2 and a printer resolution of 600 dpi, i.e. a single point had a diameter of approximately 4.23 μm .

The MTF measurements provide the image space resolution in line pairs per millimeter (lp/mm) which can be converted to object space resolution in $\mu\text{m}/\text{px}$ by

$$R_o = \frac{1000}{2 \cdot R_i \cdot m} \quad (5)$$

where R_i is the image space resolution in lp/mm determined with MTF-Mapper and m is the magnification of the camera system, i.e. the ration of sensor size versus object space FOV.

As the camera system resolution is a combination of the image space resolution and the contrast, we decided to determine the MTF at two different contrast levels. First at MTF20, i.e. a contrast of 20%, which has been deemed to be the minimum contrast for image processing algorithms, and second at MTF50, i.e. a contrast of 50%, which humans perceive as a good level. This choice is justified as images from HLIs shall help geologists to work more intuitively on data from a remote planet at MTF50. However, it should also be possible to perform further image processing on ground at

MTF20 in order to reveal even finer details in images.

Depth measurements—The accuracy of metrically calibrated plenoptic cameras has been already addressed in [6, 32]. Instead, in this experiment we focus on the depth measurement capability of a plenoptic camera in a representative setting for hand lens imaging during planetary exploration. To this end, we choose a planar surface in order to compare its measurements with the best fitted plane, and in this way assess experimentally the maximum achievable depth precision of the plenoptic camera system.

In order to measure the depth resolution and the accuracy of the calibration in depth, we created depth maps of the well textured surface of the sandblasted metal block (cf. Figure 6). Due to the 45° tilt we were able to measure depth over a wide depth range within the FOV and to observe changes of the depth precision with increasing distance to target. We expected that a tilted plane should help to reveal the depth resolution in terms of a repeating, artificial step-kind pattern in the depth data as a result from the limitations of the camera system and depth estimation algorithm. Based on the metric calibration we created metric 2-D depth maps and 3-D point clouds. For this experiment no filtering or filling operations, which are available in RxLive, were used in order to receive the raw depth measurements.

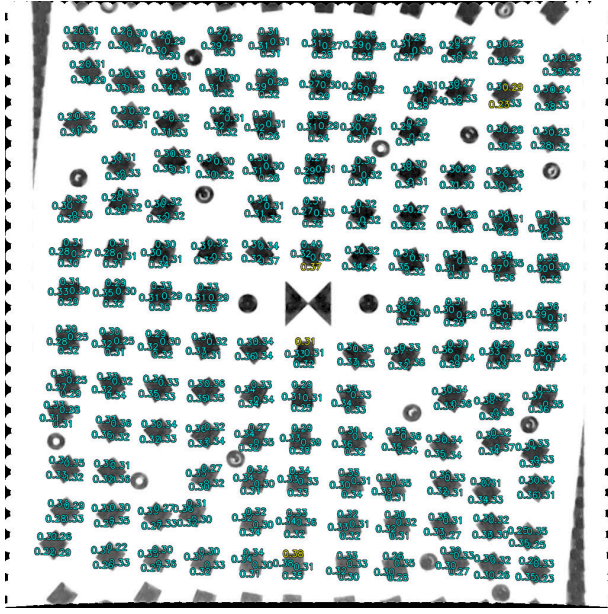
We expecteded a significant amount of depth noise in the data, which needed to be filtered previous to any analysis. In order to robustly remove outliers and segment the plane, a Random Sample Consensus (RANSAC) [38] algorithm was employed. The basic idea of RANSAC is to estimate a model fitting to the data based on a minimal, randomly chosen set of data points (in the case of a plane of three points). The algorithm then estimates the amount of inliers, i.e. points with a distance to the found model below a certain threshold and chooses again randomly a new minimal set of data points. This is iteratively repeated until a model with the maximum number of inliers is found. The advantage of this method is, that it is able to find accurate models even for very high noise levels [39]. Employing RANSAC, we segmented 63% of the raw data, the remaining inliers can be seen in Figure 9. With the filtered data, we first determined the *planarity* of the recorded data points, as well as the noise level. The planarity is calculated by the distance in millimeter of all data points with respect to the fitted plane as the Root Mean Square Error (RSME). The noise is accordingly calculated as the standard deviation (SD) in millimeter from the RMSE measurements.

Usability for planetary science—In order to provide an impression of the usability of plenoptic cameras as HLIs, we recorded data of a rock with fine structures, cavities and large depth differences. Unlike in the depth evaluation experiment, we applied depth filtering and filling operations on the data as we expect this to happen during the nominal operation as well. By this, the data becomes more easy to interpret by planetary scientists but it must be kept in mind, that some values are interpolated ones. From this data we derived an EDOF image, a depth map along with height profiles and 2.5-D models with texture and with color coded depth as shown in Section 5.

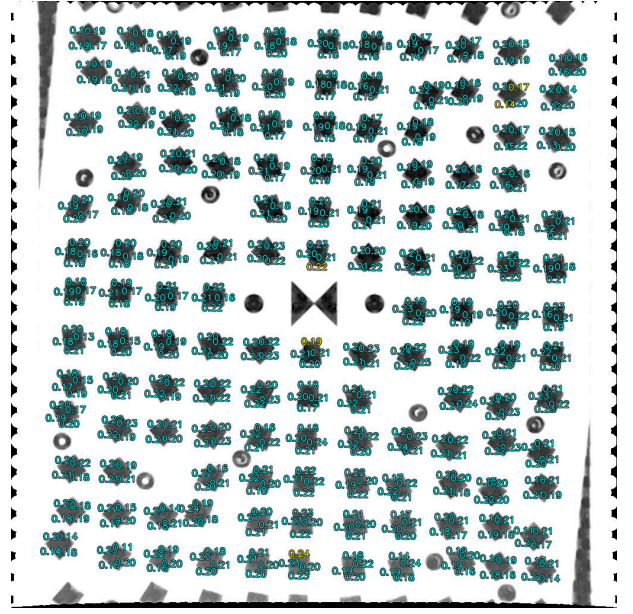
5. RESULTS

2-D resolution

Figure 7 shows the annotated lensgrid test charts for the MTF20 and the MTF50 condition, respectively. The values



(a) MTF20



(b) MTF50

Figure 7: Lensgrid test charts recorded at a WD of 160 mm with annotations for MTF20 and MTF50 values for each edge

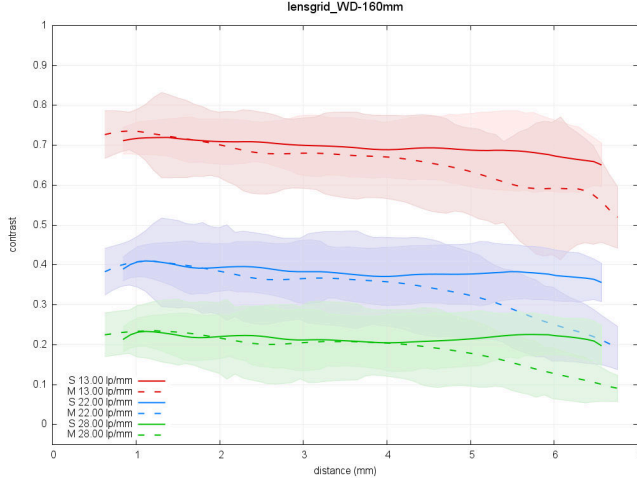


Figure 8: MTF-plot for 13 lp/mm, 22 lp/mm and 28 lp/mm computed for a lensgrid test chart recorded at a WD of 160 mm

annotating each of the edges show the image space resolution in lp/mm for which the contrast equals 20% or 50%, respectively. These plots allow to determine the resolution for the defined contrast in detail at a certain edge within the FOV. A summary of the annotated lensgrid charts is shown in the MTF profile in Figure 8. Here, the MTF values are plotted for a certain image space resolution in lp/mm as a function of the distance to the center of the FOV. This indicates how well the system maintains the resolution throughout the FOV and which contrast can be achieved at which resolution. The continuous line shows the MTF for the sagittal direction, i.e. for edges along the radial direction starting from the image center. The dotted line shows the contrast for the meridional direction, i.e. for the edges which are in tangential direction to a circle centered in the image center. The shaded parts

around the lines indicate the spread of the measurement, which can result from a misalignment between sensor plane and testchart. Because there is no increase of the spread towards the FOV border, we argue that the inherent noise of the MTF measurement becomes more visible due to the small WD. As can be seen from Figure 8, the tested plenoptic camera system maintains a nearly constant resolution over the complete FOV in sagittal direction but in the meridional direction, the contrast reduces quickly after a distance of 4.5 mm from the image center, i.e. after approximately 60% of the FOV. At the time of writing, it is not clear, yet, if this is an effect caused by the main lens or by the plenoptic data processing.

For the WD of 160 mm, we determined a FOV of 58 mm. In combination with a sensor width of 11.264 mm, this results in a magnification of $m = 0.1942$. Hence, with the camera system under test it is possible to achieve the following object space resolution according to Equation 5 and to Figure 8. Please note that in Figure 8 MTF50 relates to a contrast of 0.5. The resolution in image space given in lp/mm translate to object space as $198 \mu\text{m}/\text{px}$ for 13 lp/mm, $117 \mu\text{m}/\text{px}$ for 22 lp/mm and $92 \mu\text{m}/\text{px}$ for 28 lp/mm. Therefore, over the complete FOV either a resolution of $117 \mu\text{m}/\text{px}$ for MTF20 (blue) or of $198 \mu\text{m}/\text{px}$ for MTF50 (red) is possible. In case the observable resolution differences in sagittal and meridional direction can be handled by choosing a better main lens, a resolution of up to $92 \mu\text{m}/\text{px}$ (green, continuous) for MTF20 is possible with the system under test.

Depth measurement

Due to the observed resolution loss in 2-D, we used only 60% of the FOV for the depth analysis, i.e. the central 600 px in x- and y-direction of the depth map and the corresponding area of the resulting 3-D point cloud. Fitting the best plane to the data as explained in Section 4 results in a RMSE of $194.42 \mu\text{m}$ and a SD of $156.40 \mu\text{m}$. The visual representation of the fit is shown Figure 9. By color-coding the data points

according to their distance to the estimated plane, the error of each point becomes directly visible. Figure 10 shows the

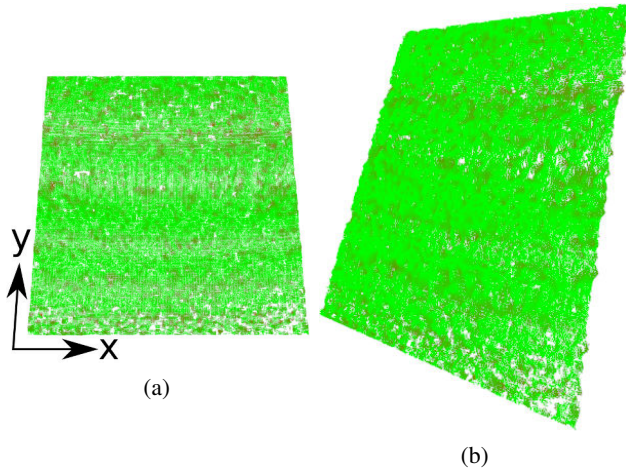


Figure 9: a) Top-down view with x-y-axis and b) perspective projection of the plane fit quality. The color coding depicts the distance of each pixel to a robustly fitted plane

average error with respect to the best fitted plane in more detail along the x- and y-direction. It can be seen, that in the y-direction the spread of the mean values is small but a certain pattern is visible in the data which is below 0.5 mm. The spread of the mean values in the x-direction is slightly larger at some positions and it does also show a pattern, but less clear than in the y-direction. These patterns are also visible as small bumps in Figure 11 or as wave-like structures in Figure 9. As Figure 10 represents a filtered and fitted point cloud, the remaining visible artefacts are most likely caused by small inaccuracies of the calibration model. Those inaccuracies are probably visible due to the sub-millimeter range of the depth precision of the presented measurements. In order to clarify the cause for this pattern a dedicated experiment is required in the future. Nevertheless, Figure 10 shows that the depth precision remains within the 0.5 mm boundary over the complete pixel dimensions, i.e. it does not change with an increasing distance between camera and object. Additionally, the slope of the mean depth values in Figure 11 shows the precise reproduction of the tilt of the metal block, which was 44.6° as measured with the metrology arm (cf. Figure 6). These two observations indicate, that the depth precision remains stable and reliable over a wide depth range. Please note, that the depth values in Figure 11 are not related to the WD but are relative to a reference plane set by the RxLive software. As the reference plane is set according to the total depth range of a measurement and therefore disconnected from the recording situation, the range of depth of nearly 30 mm and not the absolute depth values in Figure 11 is of interest.

There is less of an artificial discretization effect of the depth measurements, i.e. the aforementioned step-like pattern, visible throughout the whole depth range as it would be expected based on the description of [5]. Probably due to the slightly higher level of noise on our particular scene and camera setup conditions. Hence, the determination of the depth resolution is not possible from a tilted plane as initially planned. Therefore, a test object with a distinct depth profile, e.g. a stair with a known step size as in [32], is required for future tests. As shown, the precision of measurements evaluated by the planarity are in a sub-mm range and as long as the depth noise and the outliers can be reliably filtered,

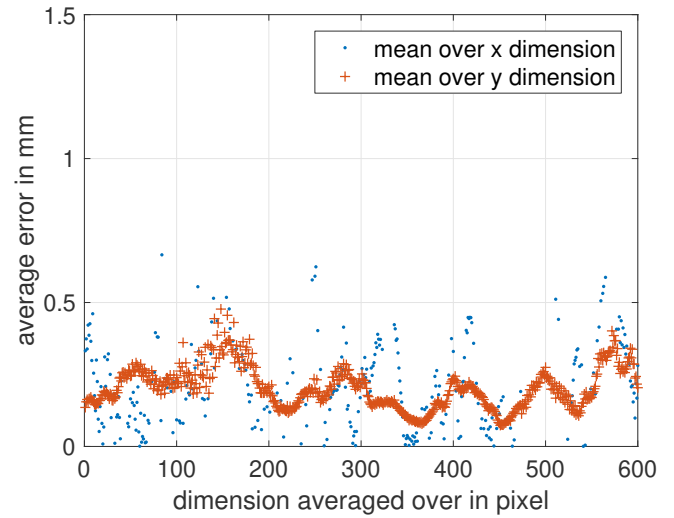


Figure 10: Average error in millimeter with respect to the fitted plane over x- and y-direction

depth differences of at least 500 μm can be recognized with our current setup.

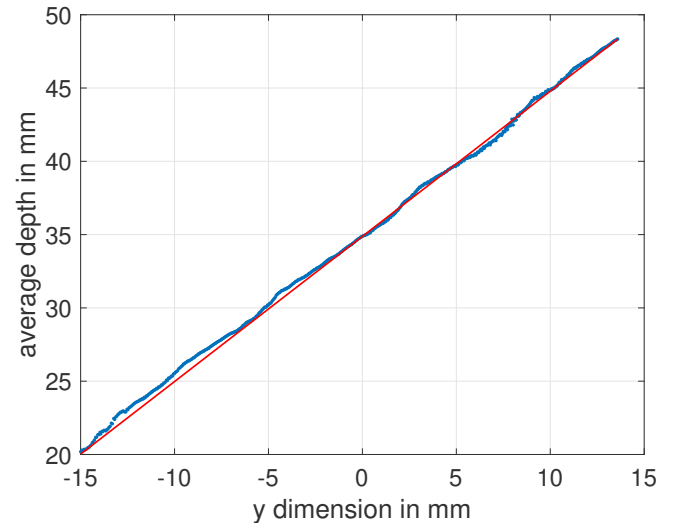
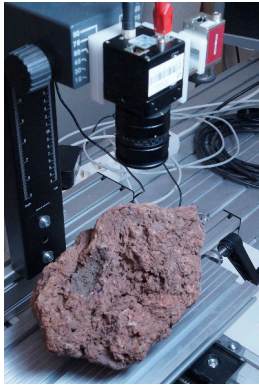


Figure 11: Average depth along y-direction in millimeter (blue) and the ground truth surface of the object plane (red)

Usage for planetology

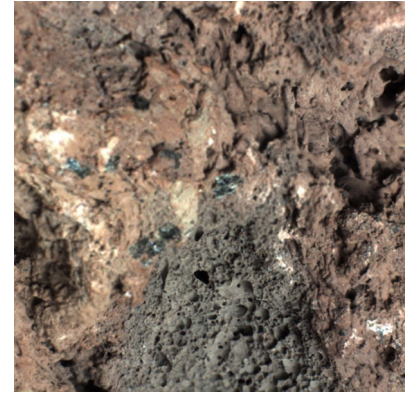
Figure 12a and 12b provide an impression of the recording situation and the context of the plenoptic data in order to ease the interpretation of the following results. The EDOF image in Figure 12c shows fine details of the rock's morphology, mineral enclosures and a pore structure on the bottom of the image. Please note, that the pore structure is approximately at the same WD as the test charts for the 2-D resolution tests, i.e. at 160 mm, whereas the remaining parts are closer to the camera. This becomes also well visible in the 2.5-D models in Figure 13, which help to interpret the depth of the scene more intuitively. As shown in Figure 13a, the model can either be textured with the EDOF image or color-coded as in Figure 13b. Augmenting the gray isolines on the model, here in intervals of 5 mm, further improves the depth impression



(a) Recording situation (based on [7])

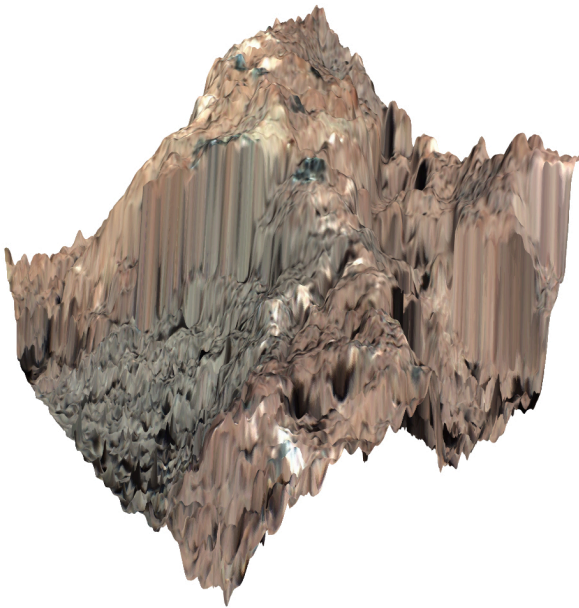


(b) Close-up showing details and depth extend of recorded area.

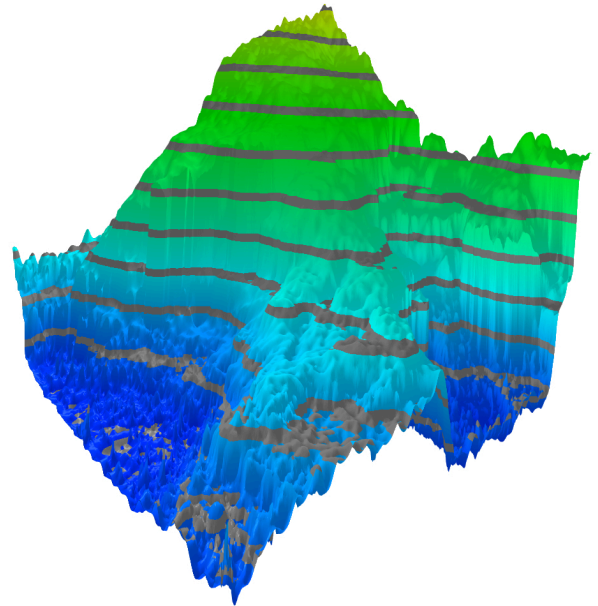


(c) EDOF image (based on [7])

Figure 12: Recording situation and context view of the target area together with the resulting EDOF image of a rock

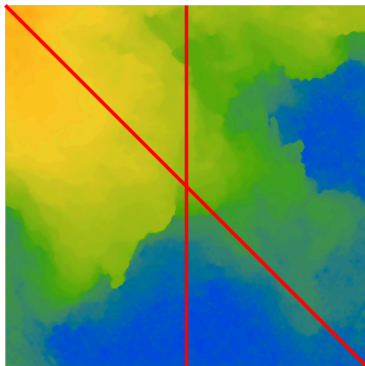


(a) Textured with EDOF image

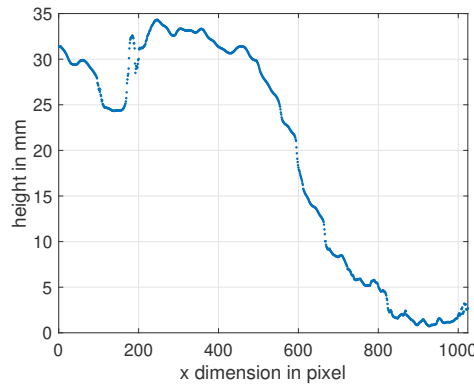


(b) Colored by depth, gray lines indicate a 5 mm height interval

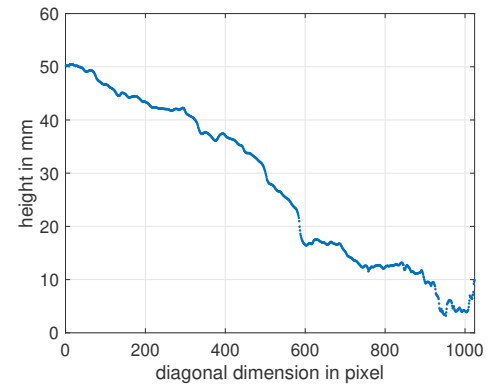
Figure 13: Renderings of a 2.5-D model of the rock with fine structures and details



(a) Colored Depth Map



(b) Height profile along the center



(c) Height profile along the diagonal

Figure 14: Color-coded depth map with indicators of the cross-cuts for the height profiles; red = near, blue = far

and the perception of the correct scale of features. During our experiments we achieved a depth range of up to 70 mm. Hence, it is possible to measure depth nearly as far as half of the WD, from the farthest point in the depth range towards the camera. By providing the 2.5-D models on a screen, a scientist can interactively explore the target, e.g. rotate or zoom or change the vantage point, which results in a more intuitive impression of the shape and nature of the target than a 2-D image can provide.

As the data is metrically calibrated, it is possible to directly pick a depth measurement of any point in the model or to create height profiles as shown in Figure 14. Here, we did a cross-cut along the center and along the diagonal as indicated by the red lines. This provides a scientist with height profiles, that e.g. allow to determine the depth of the pores visible in the EDOF image and the 2.5-D models. In Figure 14b the points in the x-axis-range between 800 and 1000 px show the pore area and given an impression of the depth resolution. With recordings taken from different vantage points, the resulting depth maps or 2.5-D models can be combined, provided the overlap between consecutive recordings is large enough. Hence, it is possible to create depth mosaics in order to provide more depth context of a scene or to create a real 3-D model of an object. In the future, it should be possible to transmit the raw plenoptic image (cf. Figure 5a) in order to gain the full control of the plenoptic processing. In case of a limited radio link bandwidth, the transmission of the smaller EDOF image and depth maps, which are processed and compressed on-board, should also be an option.

6. DISCUSSION

The HLIs for Mars, MAHLI and CLUPI, show a high performance for 2-D micro-imaging and they are valuable instruments for science, engineering, operation and public outreach. MasCam, being a pioneering instrument is designed with some different goals and priorities. Additionally, its design was constrained by the small size of the MASCOT lander and by only a few hours of operation time. The images of the descent towards the surface of asteroid Ryugu, which were available at the time of writing, are already impressive results. Despite the design differences between those instruments, we were able to compile a list of common objectives and requirements. We welcome the planetary science and the space community to discuss and enhance this list in order to be able to prepare systematically the HLIs of the future.

Our analysis shows that current HLIs are limited by the trade-off between resolution, FOV, WD and subsequently the DOF because they are based on the conventional camera concept. These physical limitations require either a complex focus mechanism, which increases the size of the instrument, risk of failure, complexity and therefore costs. Or it requires to sacrifice resolution for a more simple design as with MasCam. Furthermore, the depth measurement capabilities of these instruments are either rudimentary in case of MAHLI or CLUPI or not available as in the case of MasCam. The lessons learnt from MAHLI [18] and our discussions with DLR planetary scientists show the increasing demand and high value of precise and reliable depth measurements with high resolution.

Due to the small DOF and the required focusing, the required operation time per image can sometimes be high, which in the case of MAHLI led to less use for science than initially

intended [3]. This might also be caused by the basic autonomy of the instruments and their positioning, as the on-board processing mainly provides algorithms to create the final data products and compress them with the goal of reducing the size of data that is transmitted. And in the case of MAHLI, it is also a result of the missing single shot depth measurement capability that could help to position the HLI on the robotic arm more autonomously, with high reliability and precision in front of a target. Given the available hardware and the possible on-board processing power at the time when the instruments were designed, it is clear that a single shot depth measurement, including the on-board processing, was not possible. As technology advances, we argue for a systematic development towards a future HLI system, which can create on-line depth maps that are usable for science and for an autonomous robotic arm positioning system.

As shown in Section 3, the plenoptic camera concept delivers high quality 2-D imaging as known from current HLIs together with single shot EDOF images and depth maps, but without requiring a complex focus mechanism or movement of the camera. All of this helps for scientific, engineering and robotic tasks. It is achieved by shifting the focus operation from hardware to software, which requires more computational resource and the development of space grade plenoptic processing as it is discussed in [40]. As a plenoptic camera can be created by modifying existing camera hardware, it is possible to gain a new type of sensor without the need of a complete and expensive hardware development from scratch. What needs to be qualified in terms of hardware is the MLA and the modification procedure of an existing the camera design. As said, the real challenge in the maturation of this technology is the development of the space grade plenoptic data processing, which might also require the development of more efficient algorithms for parts of the processing chain. First results for more efficient processing were reported in literature and are for example discussed in [40].

The results presented in this work were achieved with a preliminary, non-optimized plenoptic camera system. Nevertheless, they show in practice that, with further optimization, plenoptic cameras can achieve the required 2-D ground resolution. Our system achieved in parts a ground resolution of $92 \mu\text{m}/\text{px}$. Thus, in order to identify objects of $62 \mu\text{m}$, an improvement by at least six times is required in order to achieve the required 2-D ground resolution of $15 \mu\text{m}/\text{px}$ for this use case. But in comparison with MasCam, the requirements were already achieved with an MTF20 resolution of $117 \mu\text{m}/\text{px}$ and an image resolution of 1 Mpx. By applying a higher resolution sensor, e.g. with a 12 Mpx sensor, the resulting final image size of 4 Mpx becomes equivalent to the 4.7 Mpx of CLUPI and probably the 2-D and also the depth resolution will increase significantly. In contrast to this, the depth capabilities shown in our test exceed by far the ones of HLIs that provide depth from focus stacking. The average depth error in our experiments of less than $500 \mu\text{m}$ in conjunction with a high density of the depth map allows to investigate fine details. As also discussed in our previous work [7], this increases the information content of the recording and allows the scientist a more intuitive work with the data.

7. CONCLUSION

In this work, we analyzed the current state of the art of hand lens imagers and derived a set of common goals and requirements for the development of future devices. Our

analysis shows the lack of depth measurement capabilities of hand lens imagers that are based on conventional cameras. At the same time, the need for precise and reliable depth measurements along with already available high resolution 2-D imaging is increasing. Additionally, the space grade focus mechanism in some of the current hand lens imagers makes them complex instruments. Therefore, we propose to use the plenoptic camera concept for future hand lens imagers to avoid the need for a focus mechanism while gaining equivalent 2-D imaging capabilities and a significantly increased depth measurement performance. The inherent advantages of plenoptic cameras become especially evident at close range, which is another reason why we propose this technology for in-situ close range imaging. This paper presented the first experimental performance evaluation for 2-D resolution and depth measurements for the use case of plenoptic hand lens imagers. Our results show the potential of plenoptic cameras to provide the required performance but they also indicate that in terms of optical design and especially calibration and data processing, there is still need for further development. To conclude, given the early stage of development of this technology, the performance is at least equivalent for 2-D images and significantly better for depth measurements than current conventional hand lens imagers. Therefore, our future work will consist of more detailed experiments with higher resolution plenoptic camera, the continuation of the optical modelling and the algorithm development for plenoptic hand lens imagers. Based on this, we also plan to develop more detailed concepts for space grade plenoptic hand lens imagers.

ACKNOWLEDGMENTS

We thank F. Trauthan, E. Hauber, J. P. de Vera and M. Baqué of the DLR Institute of Planetary Research for their help in deriving common goals and requirements for future HLIs. Many thanks to F. van den Bergh for the creation of MTF-Mapper and for his support during our experiments.

REFERENCES

- [1] K. Edgett, M. Ravine, M. Caplinger, F. Ghaemi, J. Schaffner, M. Malin, J. Baker, D. Dibiase, J. Laramée, J. Maki, R. Willson, J. Bell, J. Cameron, W. Dietrich, L. Edwards, B. Hallet, K. Herkenhoff, E. Heydari, L. Kah, M. Lemmon, M. Minitti, T. Olson, T. Parker, S. Rowland, J. Schieber, R. Sullivan, D. Sumner, P. Thomas, and R. Yingst, "The Mars Science Laboratory (MSL) Mars Hand Lens Imager (MAHLI) Flight Instrument," in *Lunar and Planetary Science Conference*, ser. Lunar and Planetary Inst. Technical Report, vol. 40, Mar. 2009, p. 1197.
- [2] J.-L. Josset, F. Westall, B. A. Hofmann, J. Spray, C. Cockell, S. Kempe, A. D. Griffiths, M. C. D. Sanctis, L. Colangeli, D. Koschny, K. Föllmi, E. Verrecchia, L. Diamond, M. Josset, E. J. Javaux, F. Esposito, M. Gunn, A. L. Souchon-Leitner, T. R. R. Bontognali, O. Korabiev, S. Erkmann, G. Paar, S. Ulamec, F. Foucher, P. Martin, A. Verhaeghe, M. Tanevski, and J. L. Vago, "The Close-Up Imager Onboard the ESA ExoMars Rover: Objectives, Description, Operations, and Science Validation Activities," *Astrobiology*, vol. 17, no. 6 & 7, pp. 595–611, Jun. 2017.
- [3] R. A. Yingst, K. Edgett, M. Kennedy, G. Krezoski, M. McBride, M. Minitti, M. Ravine, and R. Williams, "MAHLI on Mars: lessons learned operating a geoscience camera on a landed payload robotic arm," *Geoscientific Instrumentation, Methods and Data Systems*, vol. 5, pp. 205–217, Jun. 2016.
- [4] F. T. Ghaemi, "Design and fabrication of lenses for the color science cameras aboard the Mars Science Laboratory rover," *Optical Engineering*, vol. 48, no. 10, pp. 103 002–103 002–15, 2009.
- [5] C. Perwass and L. Wietzke, "Single lens 3D-camera with extended depth-of-field," in *Proc. SPIE 8291 Human Vision and Electronic Imaging XVII*, vol. 8291, Feb. 2012, pp. 829 108–829 108–15.
- [6] K. H. Strobl and M. Lingenauber, "Stepwise calibration of focused plenoptic cameras," *Computer Vision and Image Understanding*, vol. 145, pp. 140–147, 2016, light Field for Computer Vision.
- [7] M. Lingenauber, U. Krutz, F. A. Fröhlich, C. Nissler, and K. Strobl, "Plenoptic cameras for in-situ micro imaging," in *European Planetary Science Congress 2018*, ser. EPSC Abstracts, vol. 12. Copernicus, Sep. 2018, p. 982.
- [8] K. S. Edgett, R. A. Yingst, M. A. Ravine, M. A. Caplinger, J. N. Maki, F. T. Ghaemi, J. A. Schaffner, J. F. Bell, L. J. Edwards, K. E. Herkenhoff, E. Heydari, L. C. Kah, M. T. Lemmon, M. E. Minitti, T. S. Olson, T. J. Parker, S. K. Rowland, J. Schieber, R. J. Sullivan, D. Y. Sumner, P. C. Thomas, E. H. Jensen, J. J. Simmonds, A. J. Sengstacken, R. G. Willson, and W. Goetz, "Curiosity's Mars Hand Lens Imager (MAHLI) Investigation," *Space Science Reviews*, vol. 170, no. 1, pp. 259–317, 2012.
- [9] R. Jaumann, N. Schmitz, A. Koncz, H. Michaelis, S. E. Schroeder, S. Mottola, F. Trauthan, H. Hoffmann, T. Roatsch, D. Jobs, J. Kachlicki, B. Pforte, R. Terzer, M. Tschentscher, S. Weisse, U. Mueller, L. Perez-Prieto, B. Broll, A. Kruselburger, T.-M. Ho, J. Biele, S. Ulamec, C. Krause, M. Grott, J.-P. Bibring, S. Watanabe, S. Sugita, T. Okada, M. Yoshikawa, and H. Yabuta, "The Camera of the MASCOT Asteroid Lander on Board Hayabusa 2," *Space Science Reviews*, vol. 208, no. 1, pp. 375–400, Jul. 2017.
- [10] K. E. Herkenhoff, S. W. Squyres, J. F. Bell, J. N. Maki, H. M. Arneson, P. Bertelsen, D. I. Brown, S. A. Collins, A. Dingizian, S. T. Elliott, W. Goetz, E. C. Hagerott, A. G. Hayes, M. J. Johnson, R. L. Kirk, S. McLennan, R. V. Morris, L. M. Scherr, M. A. Schwochert, L. R. Shiraishi, G. H. Smith, L. A. Soderblom, J. N. Sohl-Dickstein, and M. V. Wadsworth, "Athena Microscopic Imager investigation," *Journal of Geophysical Research: Planets*, vol. 108, no. E12, 2003.
- [11] H. U. Keller, W. Goetz, H. Hartwig, S. F. Hviid, R. Kramm, W. J. Markiewicz, R. Reynolds, C. Shinohara, P. Smith, R. Tanner, P. Woida, R. Woida, s. B. J. Bo, and M. T. Lemmon, "Phoenix Robotic Arm Camera," *Journal of Geophysical Research: Planets*, vol. 113, no. E3, 2008.
- [12] H. U. Keller, H. Hartwig, R. Kramm, D. Koschny, W. J. Markiewicz, N. Thomas, M. Fernades, P. H. Smith, R. Reynolds, M. T. Lemmon, J. Weinberg, R. Marcialis, R. Tanner, B. J. Boss, C. Oquest, and D. A. Paige, "The MVACS Robotic Arm Camera," *Journal of Geophysical Research: Planets*, vol. 106, no. E8, pp. 17 609–17 621, Aug. 2001.
- [13] S. Mottola, G. Arnold, H.-G. Grothues, R. Jaumann, H. Michaelis, G. Neukum, and J.-P. Bibring, "The

- Rolis Experiment on the Rosetta Lander,” *Space Science Reviews*, vol. 128, no. 1, pp. 241–255, Feb. 2007.
- [14] S. Schröder, S. Mottola, G. Arnold, H.-G. Grothues, R. Jaumann, H. Keller, H. Michaelis, J.-P. Bibring, I. Pelivan, A. Koncz, K. Otto, E. Remeteau, F. Souvannavong, and B. Dolives, “Close-up images of the final Philae landing site on comet 67P/Churyumov-Gerasimenko acquired by the ROLIS camera,” *Icarus*, vol. 285, pp. 263–274, 2017.
 - [15] A. Allwood, B. Clark, D. Flannery, J. Hurowitz, L. Wade, T. Elam, M. Foote, and E. Knowles, “Texture-specific elemental analysis of rocks and soils with PIXL: The Planetary Instrument for X-ray Lithochemistry on Mars 2020,” in *2015 IEEE Aerospace Conference*, Mar. 2015, pp. 1–13.
 - [16] C. K. Wentworth, “A Scale of Grade and Class Terms for Clastic Sediments,” *The Journal of Geology*, vol. 30, no. 5, pp. 377–392, 1922.
 - [17] D. R. DiBiase and J. Laramee, “Mars hand lens imager: Lens mechanical design,” in *2009 IEEE Aerospace conference*, Mar. 2009, pp. 1–10.
 - [18] K. S. Edgett, M. A. Caplinger, J. N. Maki, M. A. R. vine, F. T. Ghaemi, S. McNair, K. E. Herkenhoff, B. M. Duston, R. G. Willson, R. A. Yingst, M. R. Kennedy, M. E. Minitti, A. J. Sengstacken, K. D. Supulver, L. J. Lipkaman, G. M. Krezoski, M. J. McBride, T. L. Jones, B. E. Nixon, J. K. V. Beek, D. J. Krysak, , and R. L. Kirk, “Curiosity’s robotic arm-mounted Mars Hand Lens Imager (MAHLI): Characterization and calibration status, MSL MAHLI Technical Report 0001,” Malin Space Science Systems, Tech. Rep., 2015.
 - [19] E. H. Adelson and J. R. Bergen, “The Plenoptic Function and the Elements of Early Vision,” *Computational Models of Visual Processing*, pp. 3–20, 1991.
 - [20] M. Levoy and P. Hanrahan, “Light Field Rendering,” in *SIGGRAPH ’96 Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH ’96. New York, NY, USA: ACM, 1996, pp. 31–42.
 - [21] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, “The lumigraph,” in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH ’96*, - 1996, p. nil.
 - [22] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, “Light field photography with a hand-held plenoptic camera,” Stanford University, Tech. Rep. 11, 2005.
 - [23] T. G. Georgiev and A. Lumsdaine, “Resolution in plenoptic cameras,” in *Computational Optical Sensing and Imaging*. Optical Society of America, 2009, p. CTuB3.
 - [24] A. Lumsdaine and T. Georgiev, “Full resolution light-field rendering,” Indiana University and Adobe Systems, Tech. Rep., 2008.
 - [25] —, “The focused plenoptic camera,” in *2009 IEEE International Conference on Computational Photography (ICCP)*, April 2009, pp. 1–8.
 - [26] T. Georgiev and A. Lumsdaine, “The multifocus plenoptic camera,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 829 908–829 908.
 - [27] H. Gross, Ed., *Handbook of Optical Systems - Fundamentals of Technical Optics*, 1st ed., ser. Handbook of Optical Systems. Wiley VCH GmbH, 2005, vol. 1.
 - [28] H. Gross, F. Blechinger, and B. Achtner, *Handbook of Optical Systems - Survey of Optical Instruments*, 1st ed., ser. Handbook of Optical Systems, H. Gross, Ed. Wiley VCH Verlag GmbH, 2008, vol. 4.
 - [29] G. Wetzstein, I. Ihrke, D. Lanman, and W. Heidrich, “Computational Plenoptic Imaging,” *Computer Graphics Forum*, vol. 30, no. 8, pp. 2397–2426, 2011.
 - [30] C. Zhou and S. K. Nayar, “Computational Cameras: Convergence of Optics and Processing,” *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3322–3340, Dec. 2011.
 - [31] U. Krutz, M. Lingenauber, K. H. Strobl, F. Fröhlich, and M. Buder, “Diffraction model of a plenoptic camera for in-situ space exploration,” in *Proceedings of the SPIE Photonics Europe 2018 Conference*, C. Fournier, M. P. Georges, and G. Popescu, Eds., vol. 10677, 2018.
 - [32] C. Heinze, S. Spyropoulos, S. Hussmann, and C. Perwass, “Automated Robust Metric Calibration Algorithm for Multifocus Plenoptic Cameras,” *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 5, pp. 1197–1205, May 2016.
 - [33] N. Zeller, F. Quint, and U. Stilla, “Calibration and accuracy analysis of a focused plenoptic camera,” *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, pp. 205–212, 2014.
 - [34] —, “Depth estimation and camera calibration of a focused plenoptic camera for visual odometry,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 118, pp. 83–100, 2016.
 - [35] T. Georgiev and A. Lumsdaine, “Focused plenoptic camera and rendering,” *Journal of Electronic Imaging*, vol. 19, no. 2, pp. 021 106–021 106, 2010.
 - [36] F. van den Bergh. (2018) MTF mapper. Web. Last accessed 2018-10-19. [Online]. Available: <https://sourceforge.net/projects/mtfmapper/>
 - [37] P. D. Burns *et al.*, “Slanted-edge MTF for digital camera and scanner analysis,” in *Is and Ts Pics Conference*. SOCIETY FOR IMAGING SCIENCE & TECHNOLOGY, 2000, pp. 135–138.
 - [38] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
 - [39] C. Nissler, Z.-C. Marton, and M. Suppa, “Sample consensus fitting of bivariate polynomials for initializing EM-based modeling of smooth 3D surfaces,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 4228–4234.
 - [40] M. Lingenauber, K. H. Strobl, S. Kriegel, and N. W. Oumer, “Benefits of Plenoptic Cameras for Robot Vision During Close Range On-Orbit Servicing Maneuvers,” in *Proceedings of the IEEE Aerospace Conference 2017*. IEEE, Mar. 2017, pp. 1–18.

BIOGRAPHY



Martin Lingenauber is a research scientist at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. In 2010 he graduated with distinction as a Dipl.-Ing. in aerospace engineering from the Technical University of Berlin. From 2010 until 2011 he was with the control systems department of the European Space Agency (ESA) in Noordwijk, Netherlands where he worked on testing components for attitude control systems and on new vision sensor concepts. Since 2011 he is with the Institute of Robotics and Mechatronics and works on robot vision for space applications and plenoptic cameras for robotics.



Klaus H. Strobl is a group leader at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. His research interests focus on computer vision, 3-D graphics, camera calibration, mobile robotics, and deep learning. Klaus studied electrical engineering (automatic control) at the Universidad de Navarra (Spain), the Vienna University of Technology (Austria), the Technische Universität München (Germany), and the Norwegian University of Science and Technology (Norway). He held a visiting researcher position at the Department of Computing, Imperial College London in 2009 and earned his Ph.D. *summa cum laude* in electrical engineering in 2014 at Technische Universität München.



Ulrike Krutz is a research scientist at the Institute of Optical Sensor Systems of the German Aerospace Center (DLR) in Berlin, Germany. She received her Dipl.-Math. degree in mathematics in 2007 at the University Essen and her Ph.D. in computer science in 2011 from the Humboldt University in Berlin. Since 2007 she has been working at the DLR at different space- and earthborn research projects with special focus on optical system theory.



Florian A. Fröhlich Florian A. Fröhlich is a research scientist at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. He studied Technical Cybernetics at the University of Stuttgart and graduated in 2008 as a Dipl.-Ing. His research interests include computer vision, 3D visualization and augmentation, tracking and segmentation.



Christian Nissler is a research scientist at the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. He studied Technical Cybernetics at the University of Stuttgart and the Technical University of Helsinki and graduated in 2011 as a Dipl.-Ing. His research interests include computer vision, localization, pose estimation, calibration and machine learning. Since 2011 Christian is working at the Department of Perception and Cognition at the Institute of Robotics and Mechatronics and was involved in several projects, focusing mainly on camera based pose estimation of robotic systems.